

Paper 2—Option C: Web Science**SL/HL Core**

C1 A publishing company, ABC Publications, based in London has a large IT department. This department is responsible for

- providing IT services to the company
- maintaining the company's web site
- creating and maintaining web based learning resources that are sold to schools and colleges

The company is finding it difficult to recruit and retain sufficient high quality IT staff to keep these functions operating at an optimal level. It is investigating transferring at least part of its IT operations to a cloud computing solution. At the moment it has not been decided how much of this should be implemented by a private cloud and how much by a public cloud.

a) Define the term private cloud.

[1 mark]

Cloud computing services are provided for a particular group with a limited number of users.

b) Distinguish between a cloud computing model and a conventional client server model in providing computing services.

[4 marks]

Traditionally with a client-server model, the client computer interacts with servers owned by the organisation itself. These servers may be located on a LAN in the offices of the organisation and be maintained by the IT team. With cloud computing, organisations outsource tasks such as storage, maintenance and data processing to outside companies, using the internet to connect client to server and allowing access from any location with an internet connection.

c) Explain why ABC Publications might benefit from changing part of its IT provision to a cloud based model.

[4 marks]

The company would no longer have to struggle to recruit and retain high quality IT staff as services would be managed by the external company. The costs associated with this as well as other jobs required to site servers locally (including maintenance, depreciation and upgrading) would be removed, and instead the company could possibly have one fixed price to pay based on their usage. In addition to this, storing information off site (potentially across many) reduces the severity of damages (fires etc.) at the local company's base. ABC might also benefit from the external companies expertise as it may have more employees dedicated to resolving network issues than if they were in ABC's office only.

d) Comment on the privacy **and** security issues relating to ABC's use of cloud computing. [4 marks]

Privacy: The use of an outside company means that any sensitive information will de facto be exposed to people outside of the ABC company. This private company may even outsource to another company so increasing the exposure of their data.

Security: If the "Cloud Company" is used to store data then there will be inevitable concerns about the reliability of this type of storage which is now out their (ABC's) hands. Cloud computing is a relatively new concept, but already companies set up to store massive amounts of (other people's) data have gone under (taking the data with them).

C2 A theatre box office maintains a web site that can display what productions are coming up and which seats are available for a particular production. It can then take orders online. The interface consists of dynamic web pages such as the one below, in which the underlying HTML interacts with client-side and server-side scripts.

BOX OFFICE

Quick search

From

To

Event type

Search

Only those performances and events currently bookable will appear in the calendar.

Visit our **What's On** page to see full production listings.

Browse

July 2015

Event type

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
11 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>	12 10:00 - 12:00 Ballet workshop <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div> 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>	13 10:00 - 12:00 Ballet workshop <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div> 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>	14 10:00 - 12:00 Ballet workshop <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div> 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>	15 10:00 - 12:00 Ballet workshop <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div> 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>	16 10:00 - 12:00 Ballet workshop <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div> 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>	17 19:30 - 22:00 <u>Swan Lake</u> 3 Counties Ballet <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin: 0 auto;">Buy now</div>

Part of the source code for this page is:

```
<script type="text/javascript" src="http://assets.ophse.org/inc/popdt/init_live.js"></script>
<script type="text/javascript" src="http://assets.ophse.org/inc/popdt/src/flash.js"></script>
```

a) Identify **one** characteristic of Hypertext Markup Language (HTML).

[1 mark]

Can be rendered by all internet browsers.

b) Identify the steps that the server would carry out so that the information in the events' calendar can be displayed on the client's computer. (4 marks)

- server retrieves event for days to be displayed from a database server
- takes results and generates HTML to display them in a table
- embeds HTML in page
- page sent to browser

c) (i) Identify two ways that a client side script may be made available to a web browser [2 marks]

A client side script may be made available to a browser by the use of script tags that are embedded in the HTML code.

By the use of an external file.

(ii) Describe one reason why a client-side script may be used in preference to a server side script. [2 marks]

A client side script will not require access to a remote server so that any processing that is done will be done more quickly and use less bandwidth. This will reduce the load on the server.

d) The organisers of the theatre want to ensure their web pages appear higher up the ranking of search engines.

Suggest whether the use of meta-tags can help achieve this aim. [4 marks]

Meta-tags are a means of promoting the key components of a website. Using these tags allows a Search Engine robot to easily pick up on the website's content, and may result in the web page ranking higher on the their website. As search engines work in a number of ways, it is difficult to quantify the importance of meta-tags, with some websites deeming the number of links towards a page more important.

e) It is common for dynamic web pages to make use of a mixture of client-side and server-side scripting. Explain how the interaction of HTML, client-side and server-side scripting has allowed the production of a web page such as the one shown here. [6 marks]

For a dynamic website such as this, prior to the HTML being generated, scripts on the server-side must be run. These may involve using a language such as PHP to handle user inputs, retrieve information from databases stored on servers, and perform calculations. As the actions of server-side calculations cannot be seen by the client, it allows for private/sensitive information to be handled

without worrying about data insecurities. On the server-side, this information will be collated and sent to the browser. The information sent may be in the form of HTML, but could take any form. A webpage could for example, output HTML with inline JavaScript, which would then be interpreted by the browser on the client side. The public nature of the client-side means it is most adapted to perform simple tasks, such as input validation/verification, but can be partnered with Server-Side scripting (e.g. through Ajax) to deliver server-driven information on the fly.

C3 As the web has developed, data formats, communication protocols and standards such as XML or SQL have proven crucial to progress. Two fundamental concerns that have been central to this development have been the issues of interoperability and that of open standards.

a) (i) Identify one characteristic of XML.

[1 mark]

It does not contain a fixed set of tags, therefore new ones can be added.

(ii) Define the term protocol.

[1 mark]

A set of rules and procedures that both sender and receiver must adhere to in order to allow coherent data transfer.

b) Describe, with the use of examples, how the use of open standards allows interoperability to occur.

[3 marks]

The use of open standards implies that anyone can use them, and standards can be introduced by anyone and agreed upon to ensure quality is maintained. Open standards require that interoperability is decided from the beginning (ab-initio), and used throughout. For example, the Internet backbone relies on the IP protocol which is an agreed standard, allowing the transfer of information to occur.

Music is distributed across the web in a variety of different ways such as peer-2-peer (P2P) networks.

c) Discuss two factors that would affect the decision to use either lossless or lossy compression when transferring files across the Internet.

[6 marks]

Lossless compression is used when no loss of data is wanted when transferring files such as audio files. However there may be cases when the data that may be lost using lossy compression will not significantly affect the final version of the file when it is decompressed, so in these cases there may be no need to use lossless compression. The decision to use either lossy or lossless compression may be more affected by the needs of the end-user who may wish to use a format such as MP3 that allows more music to be stored on a device.

Additionally the additional time taken to transmit the losslessly compressed file, which may be three or four times the size of the lossy compression, may be an important factor in choosing the compression type. For example, if there is either a fast internet connection or the file does not need to be accessed immediately then lossless compression can be used. However, with increasingly higher traffic across the Internet lossless compression may prove to be too slow.

d) Explain **one** advantage of the use of a peer-2-peer (P2P) network for obtaining and downloading music and movie files. [2 marks]

One advantage of using a P2P network rather than a client-server network for obtaining and downloading files is that it is easier to set up and less time will need to be spent in configuring the network.

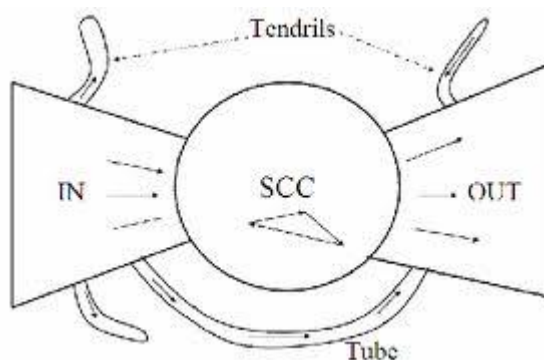
HL Extension

C4 The World Wide Web (web) can be regarded as a directed graph. This allows search engines to make use of algorithms based on graph theory.

a) Identify how the web may be represented as a directed graph. [1 mark]

The web is represented as a directed graph where the web-pages are seen as vertices and the hyperlinks between them as edges.

The web can be represented as having a bowtie structure as indicated in the diagram below.



nlp.stanford.edu. Accessed 13 April 2011

b) (i) Define the term Strongly Connected Core(SCC) [1 mark]

The strongly connected core is the part of the web where a web surfer is able to navigate to and from any particular web page.

(ii) Outline the characteristics of web sites that are located in the IN portion of the bowtie [2 marks]

These web sites tend to have a number of links that lead from them to other web sites, such as those in the SCC or the OUT portion. These pages may be seen as undiscovered destinations that have not yet been found by pages in the SCC or they may be a resource list that provides a gateway to pages in the SCC.

With the growth of the web, web site developers have realised that there are possible concerns about the ability of being able to link to all web pages as well as ensuring a page is highly ranked by search engines.

c) Describe how power laws suggest it will be possible to link from one web page to any other web page despite the fact the web is growing so rapidly. [2 marks]

Power laws say that the diameter of the web is related to the size of the web in an exponential manner. This means that if the number of web pages in the SCC increases by 10 times, the diameter of the SCC may only increase by three.

d) Explain how the relative importance of a web page can be determined in search engines. [4 marks]

The monitoring the number of links into a website is very important for a number of Search Engines such as with Google, with the most popular usually being assigned a greater page rank and being displayed further up the results page.

The rank is determined by the number of votes cast for it. This is based on the number of 'in' links as well as the importance of the pages voting for it. To push the website higher up the page rankings webmasters will try to maximise the number of 'in' links. However, some ranking algorithms may ignore some of the links as they may be generated by link farms and are not believed to be genuine, also older pages may be advantaged as they will have had time to acquire in links.

The development of the web has changes the way that users interact with the web and with each other.

e) With reference to specific examples, distinguish between an ontology and a folksonomy. [2 marks]

Ontologies are a formal framework for the storing of knowledge and the relationship between the entities in a certain field of knowledge, for example linked to a particular make of car. A folksonomy is an informal framework developed from the creation and management of tags to categorise information for example on sites like Flickr.

The growth of the web has enabled the development of new ways to solve problems. Collective intelligence is one such approach.

f) Identify **two** characteristics of collective intelligence [2 marks]

- May refer to the group mind
- Involves groups of people working together across the world using open data in a public environment.

g) With the increasing world population and the effects of globalisation, the world's population is facing new and complex problems such as the recent banking crisis. To what extent is it likely that collective intelligence could help to address these problems? [6 marks]

Collective intelligence certainly has the potential to address problems such as the banking crisis, where ignorance, causing major incorrect decisions made only by a select few, has burdened the majority of the public. With collective intelligence, anyone can have an input into the production process, and with so many being able to view and critique the work, issues are likely to be quickly discovered and resolved which may not be possible using traditional techniques such as face to face meetings. The high number of people involved in collective intelligence projects will likely include experts from a variety of different fields who can overlook specific areas of the design – this would not be possible or cost effective in a private project. However, in reality collective intelligence may struggle due to either the lack of motivation, financial or otherwise, from the public to get involved or from too much information being provided from this collaboration and it not being easy to manage or rely on systems such as voting forums to make decisions. Overall, the possible issues caused by possibly too much information, too many contributors and their lack of motivation are outweighed by the ability to gather and process large amounts of information rapidly to resolve complex problems.